PM-566 Final Project
November 23, 2020
Tao Huang

Research on the relationship between Covid-19 and Crime in LA

Introduction

Background:

In 2020, the covid-19 epidemic has swept the world, and the United States is no exception. The impact of this epidemic is all-round, whether economy, politics, diplomacy, environment, or even urban security have all been affected a lot. Los Angeles is the second largest city in the United States and the largest city in California. It is known for its ethnic diversity and economic development. Therefore, I think Los Angeles is a representative American city, and decided to choose Los Angeles as a key research object to explore the urban security situation under the covid-19.

API and Data:

This research involves two data sets. The first data set is the Los Angeles crime record data set. Thanks to lacity.org for providing my data API, allowing me to obtain all criminal records in Los Angeles from January 2017 to October 2020. It includes some variables that are extremely important to this research, such as the time of the crime happened, the latitude and longitude, and the type of crime. (This data set is denoted as "LA_raw_crime", Data link: https://data.lacity.org/resource/2nrs-mtv8.json)

The second data set is a summary of the Los Angeles covid-19 report, from the Los Angeles Times. The data set contains the daily number of new covid-19 patients with covid-19 in the Los Angeles area from March 10, 2020 to October 7, 2020. (This data set is denoted as "LA_raw_covid", Data link: https://raw.githubusercontent.com/Icygrey/pm566_lab12_hw5/main/latimes-place-totals.csv)

Methods

Data Preparation：

First of all, as the most basic and most important step in the data analysis step — data cleaning, I clean and organize the data of the LA_raw_crime data set, which involves some more cumbersome steps:

1. standardize the format of the time variable.
2. Extract the latitude and longitude and perform approximate processing.
3. Calculate the daily crime frequency and cumulative crime frequency.
4. Use regular expressions to extract the type of crime and calculate the frequency. Similarly, our LA_raw_covid data set performs similar processing.

Process：

Then, I tried to simultaneously display four time series curves on one plot to compare the difference in cumulative crime frequency from 2017 to 2020. Considering that interactive charts are easier to express information, here I use the "Plotly" package to draw, and set pop-ups at each point to display the frequency of the corresponding date. (See figure 1)

Next, I used the "Leaflet" package to draw two Head maps of the Los Angeles area. The first one is the Head maps of the crime frequency, and the second one is the Head maps of the positive cases of covid-19. I use gradation of color to indicate the severity of different areas. The darker red area indicates the greater the frequency. In order to facilitate and intuitively obtain data, I set up a pop-up in each area to display the frequency of covid-19 positive patients in the corresponding area. (See figure 2, 3)

Finally, I drew three bar graphs to compare the changes in crime types in 2020 and 2019. In the first and second plots, I directly draw the bar graphs according to the frequency of crime types in 2019 and 2020 in descending order, so that we can observe which crime types occur more frequently. (See figures 4 and 5). In order to represent the third bar graph well, I redesigned the data set behind it. I used proportion instead of frequency as the y-axis, and used overlapping bar graphs to compare and show the rise and fall of the proportion of various crimes in 2019 and 2020. (See figure 6)

Results

My final website link: https://icygrey.github.io/pm566_lab12_hw5/

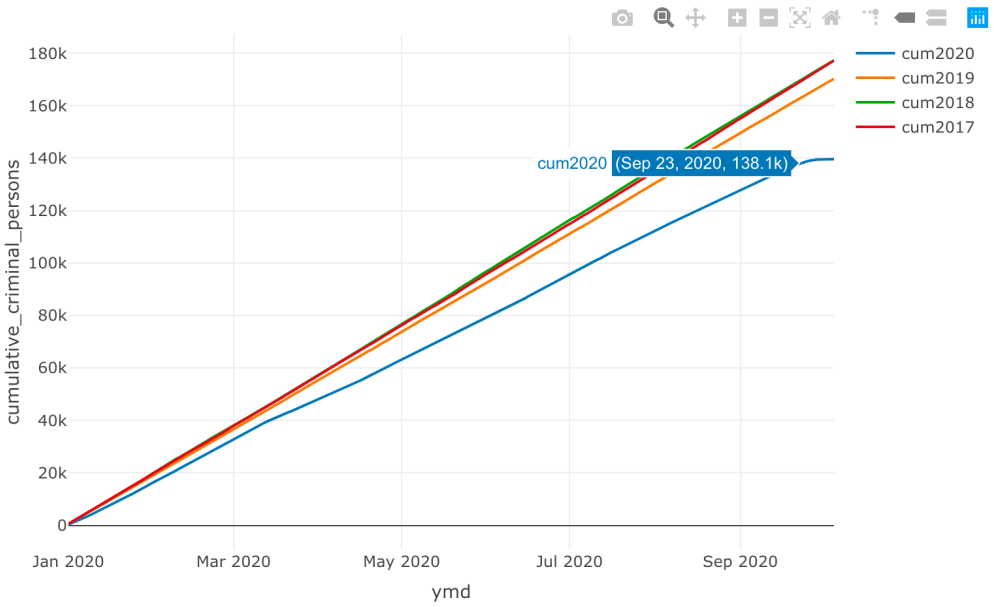

Figure 1: Frequency of cumulative crime in LA

By comparing the frequency of cumulative crime in LA for 4 years (2017-2020), it can be clearly found that the frequency of cumulative crime in LA in 2020 (blue line) has decreased significantly, especially after the covid-19 outbreak. (From February to March, the blue line starts to deviate from the other lines.)

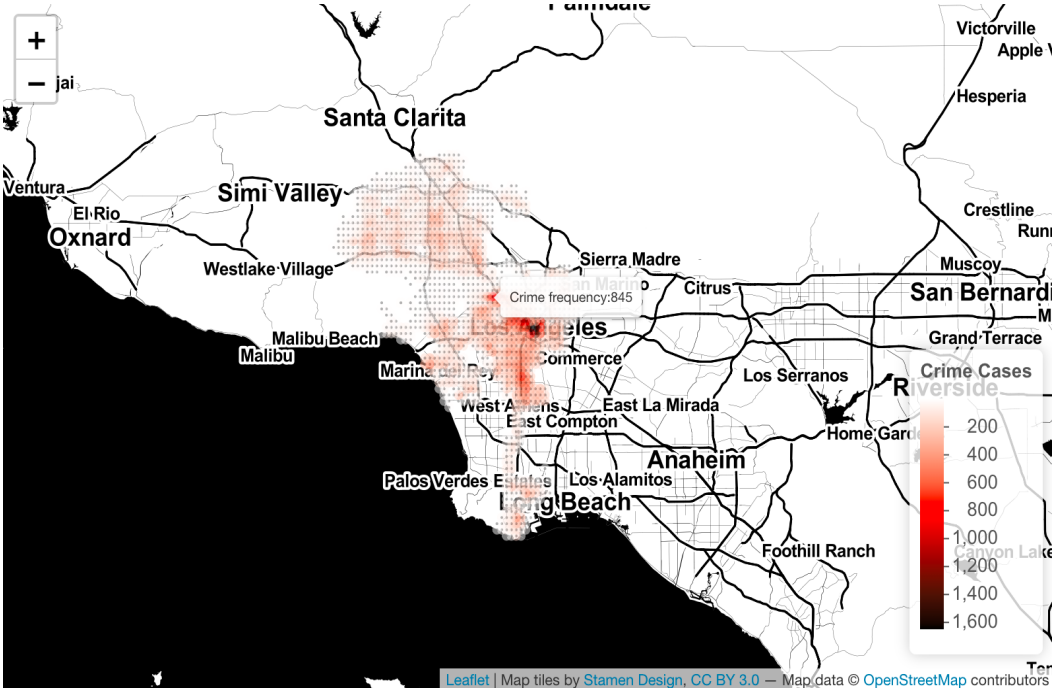

Figure 2: Crime Heat Map in LA

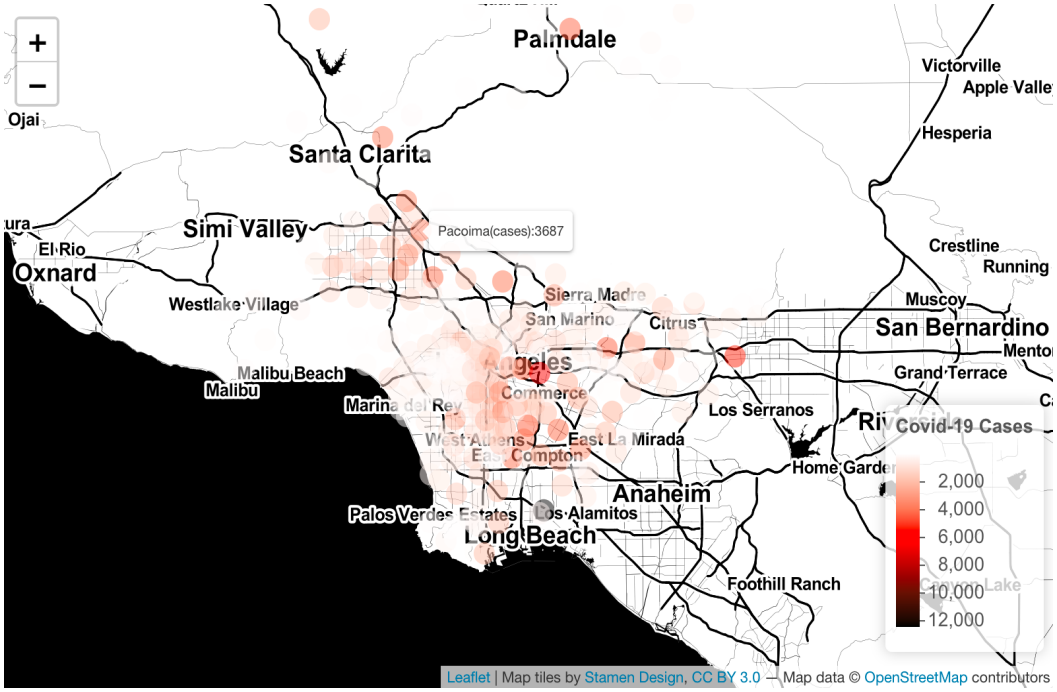

Figure 3: Covid-19 confirmed population Heat Map in LA

Through the comparison of the two Heat Maps, we can find that downtown and central-southern Los Angeles are more likely to become covid-19 outbreak points and have a higher frequency of crime in 2020.(See the dark red and black areas in the map)

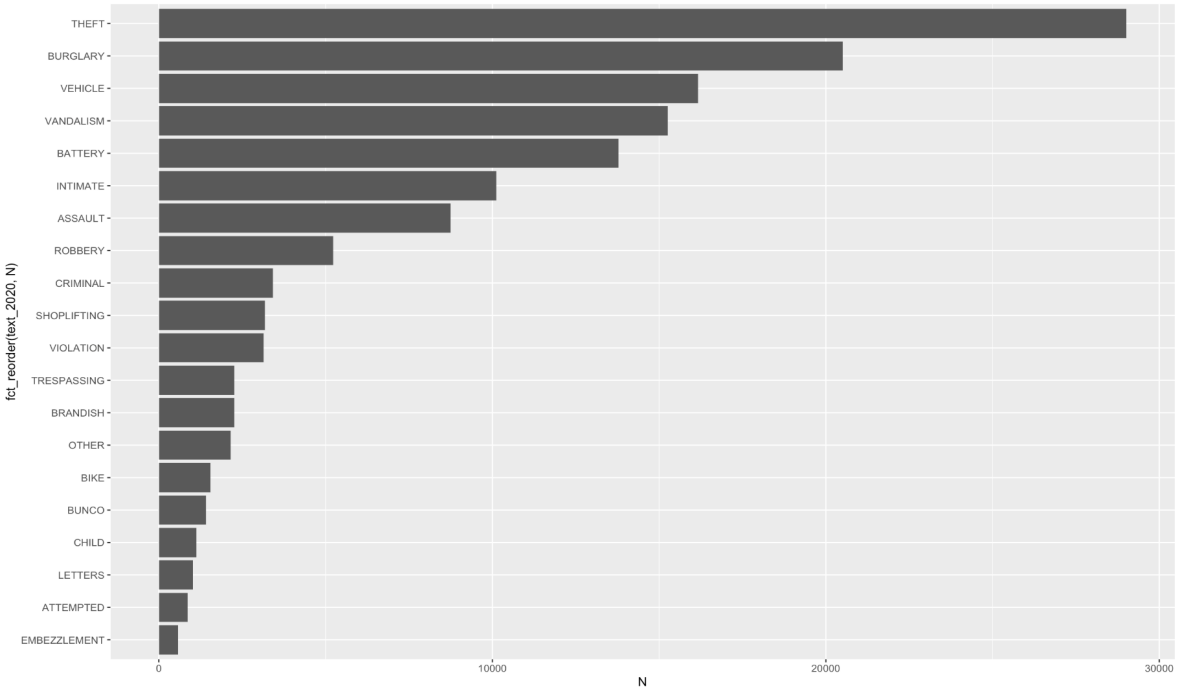

Figure 4: Types of crime occurred ranking in 2020

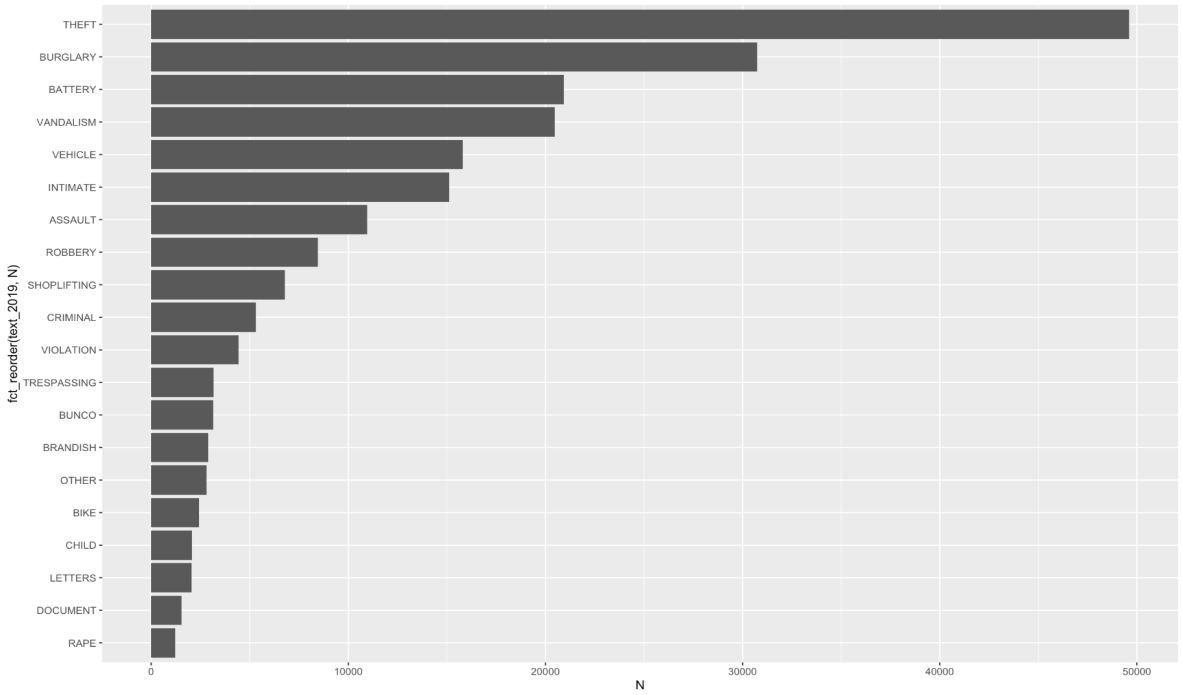

Figure 5: Types of crime occurred ranking in 2019

From the figures above, It can be found that the frequency of economic crimes (Theft, Burglary) in 2019 and 2020 both occupy the first and second places.

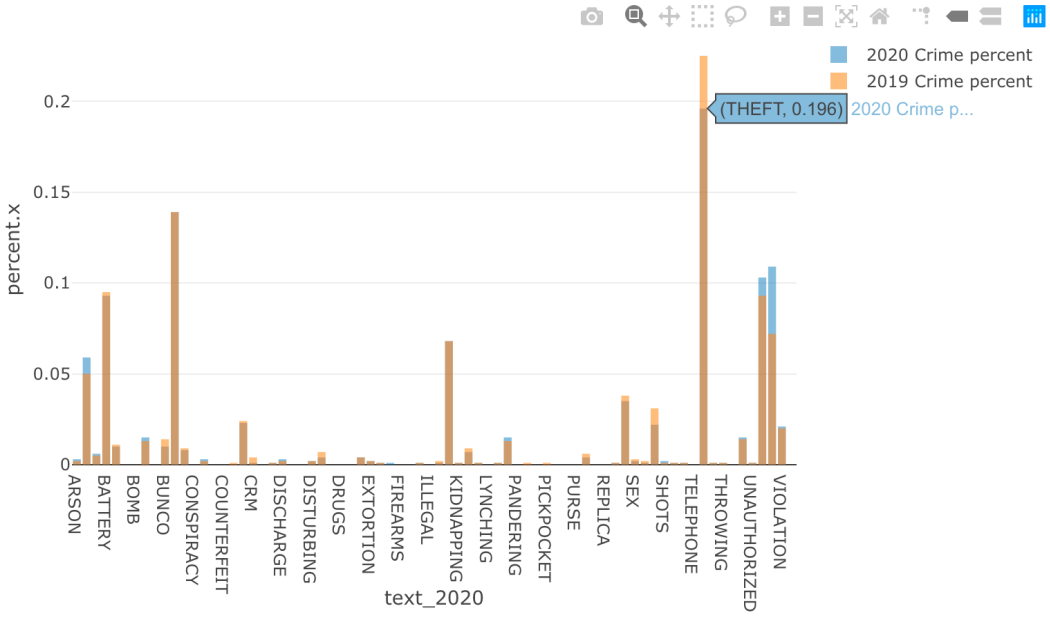

Figure 6: Compare the frequency of various crimes in 2019 and 2020

In 2020, Los Angeles economic crimes (Theft and Burglary) accounted for a decrease in the proportion of total crimes in the year compared with 2019.

In addition, we can clearly see the decrease in the frequency and proportion of battery crime, and the frequency of occurrence has dropped from the third place in 2019 to the fifth place in 2020. On the contrary, the frequency of vehicle occurrence has risen to third place in 2020. This is consistent with our impression that during the epidemic, people prefer to use private cars to travel instead of taking public transportation, which may indirectly increase the number of traffic accidents.

Conclusion

The crime frequency in Los Angeles in 2020 has been significantly reduced since February compared to previous years. I guess, this may be because the outbreak of covid-19 has allowed more residents to choose not to go out, thus reducing the risk of harm. In addition, through the comparison of heat maps, communities with densely populated (downtown) and low-income groups (central-southern Los Angeles) are more likely to become covid-19 outbreak points and have a higher frequency of crime.

Besides, through text mining, we found that crimes in the past two years were concentrated in economic cases (Theft and Burglary), and In 2020, Los Angeles economic crimes accounted for a decrease in the proportion of total crimes in the year compared with 2019. I guess that this may be due to the large-scale shops closed, and people reduced their outings, thus avoiding theft and robbery. And Federal government relief payments may also have played an important role.